The Art of Investing: The Secrets of Gold and Bitcoin Prices

Zhuofan Zhong, Shijie Gao, Jiahui Huang, Haoyu Zhou, Siyu Xu, Yang Lin

Hangzhou Normal University , Hangzhou 310000, China.

Abstract: Market traders frequently buy and sell volatile assets with the goal of maximizing total returns. In the complex market, how to predict the general trend of the market more accurately, how to determine the buying point or selling point, to maximize the target income, is the primary consideration of investors more scientifically and reasonably. In this paper, based on the closing price data of bitcoin and gold on Nasdaq from September 11, 2016 to September 10, 2021,[1] we first build a LSTM neural network prediction model for sliding sequence prediction, on this basis, we build a trading strategy selection model based on nonlinear programming, and introduce Monte Carlo algorithm to optimize the solution.

Keywords: LSTM Model；Sharp Ratio；Nonlinear Programming；Monte Carlo Algorithm.

1. Restatement of the Problem

We will start at \$1000 on September 11, 2016, and use the five-year trading period from September 11, 2016 to September 10, 2021. On each trading day, we will have a portfolio of cash, gold and bitcoin made up of dollars, troy ounces, and bitcoin. The initial state is [1000,0,0]. The commission for each transaction (purchase or sale) is α% of the transaction amount. Now assume $α_{gold} = 1\%$ and $α_{bitcoin} = 2\%$. There is no cost to hold an asset. And we should only use these data to solve the following problems:

- Develop a model that gives the best daily trading strategy based only on price data up to that day. Using the model and strategy, calculate the return on the initial \$1000 investment on September 10, 2021.

2. Model I: Time series prediction based on LSTM

2.1 Establishment and solution of the model

In this experiment, Model model is used to build a five-layer LSTM network model to analyze and predict gold and bitcoin respectively. In terms of data application, we take 70% of the closing price data before the forecast date as the training set for training, and 30% of the data blocks after the original data are taken for sliding sequence prediction.[2]

In the adjustment stage of model parameters: in the process of model training, we constantly adjust the step size, the number of iterations and the value of regularized parameters in the LSTM layer until the prediction effect is the best[3]. The final network parameters of the model are designed as follows:

- Time steps: 15;
- Number of hidden layers:2;
- Number of iterations:20.

According to the above steps, we get the pricing data of Bitcoin and gold as shown in the following figure：
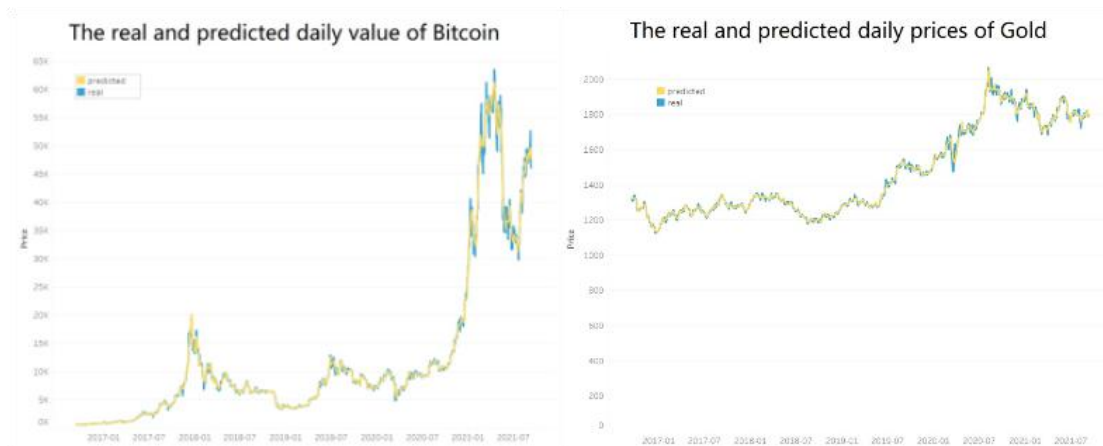
(a): Pricing of Bitcoin   (b): Pricing of Gold

Figure 3-1: Comparison of real and predicted values of Bitcoin and gold

2.2 Accuracy Test of LSTM Model

After the prediction is completed, the evaluation index of the prediction performance of the model uses Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) to compare the experimental results. The calculation formulas of RMSE and MAE are as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} \left(X_{prediction,t} - X_{real,t}\right)^2} \tag{1}$$

$$MAPE = \frac{100\%}{N} \sum_{i=1}^{N} \left|\frac{X_{prediction,i} - X_{real,i}}{X_{real,i}}\right| \tag{2}$$

Note: where N represents the number of days with forecast data, $X_{prediction,t}$ represents the forecast data of day t, and $X_{real,t}$ represents the real value of the day t.

According to the above formula, we test the accuracy of the prediction data of gold and bitcoin, and get the following results: [4]

Table1: Accuracy Test of Bitcoin

Index	RMSE	MAPE	Accurancy
Value	749.417	0.025	97.5%

Table2: Accuracy Test of Gold

Index	RMSE	MAPE	Accurancy
Value	12.952	0.005	99.5%

According to the data in the above table, we can see that the prediction accuracy of asset price trend based on LSTM neural network model is high, reaching 97.5% and 99.5% respectively, so our prediction model is more reliable, which lays a solid foundation for the establishment of investment strategy later.

3. Model II: Transaction Strategy based on nonlinear programming

3.1 The introduction of Sharp ratio

Our team introduces the concept of Sharp ratio, which is a comprehensive indicator that can consider both income and risk at the same time, to eliminate the adverse impact of risk factors on performance evaluation.[5]

The formula for calculating the Sharpe ratio is as follows:

$$Sharpe\ Ratio = \frac{E(R_p) - R_f}{\sigma_p} \tag{3}$$

Note: here $E(R_p)$ is the expected rate of return on investment, R_f is the risk-free interest rate, and σ_p is the standard deviation of investment.

$E(R_p)$ and σ_p can be quantified by the following formula：

$$E(R_p) = \frac{Expected\ return}{Investment\ cost} \times 100\% \tag{4}$$

$$\sigma_p = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}} \tag{5}$$

Note: Where x_i is the i-day rate of return, $\bar{x}$ is the average rate of return, and n is the number of days.

3.2 Establishment of nonlinear programming Model

Here are the symbols we will use in future discussions：

Table3: Notations we use in future discussion

Symbol	Description
n	Cash holdings before the start of the transaction
n_g	Gold holdings before start of the transaction
n_b	Bitcoin holdings before the start of the transaction
x_g	Gold trading volume on the trading day
x_b	Bitcoin trading volume on the trading day
$P_{g,i}, i = 0,1,2,3$	The price of gold on the i day after the trading day
$P_{b,i}, i = 0,1,2,3$	The price of bitcoin on the i day after the trading day
$I_i, i = 1,2,3$	Expected rate of return on the i day after the trading day
α_g	Commission on each transaction (gold)
α_b	Commission on each transaction (bitcoin)

Since the higher the Sharpe ratio, the higher the excess rate of return per unit risk of the trading strategy, the better the performance, we use Sharpe ratio as a measure to establish a nonlinear programming model.

In the following, we calculate the Sharp ratio every three days. In addition, in the model assumption, we have stated that there are no costs and returns for cash holdings, so $R_f = 0$. Based on this, we can get the following objective function:

$$Max\ sharpe\ ratio = \frac{E(R_p)}{\sigma_p} \tag{6}$$

Compared with the day of trading, the rate of return of the day after trading is as follows：

$$I_1 = \frac{P_{g,0}|x_g|\alpha_g\% - P_{b,0}|x_b|\alpha_b\% + (n_g + x_g)(P_{g,1} - P_{g,0}) + (n_b + x_b)(P_{b,1} - P_{b,0})}{n + P_{g,0}n_g + P_{b,0}n_b}$$

In the same way, the expressions of I_2 and I_3 can be obtained. Finally, the expression of the objective function of our simultaneous equation is as follows：

$$Max\ sharpe\ ratio = \frac{I_1 + I_2 + I_3}{\sqrt{3\sum_{i=1}^{3}(I_i - \bar{I})^2}} \tag{7}$$

After determining the objective function, we give the following constraints of the programming model:

$$\begin{cases} P_{g,0}x_g + P_{b,0}x_b + P_{g,0}\left|x_g\right|\alpha_g\% + P_{b,0}\left|x_b\right|\alpha_b\% \leqslant n \\ -x_g \leq n_g \\ -x_b \leq n_b \\ x_g = 0 \qquad If\ the\ trading\ day\ is\ a\ weekend \end{cases}$$

3.3 Establishment of nonlinear programming Model

For the above programming model, it is a deterministic problem, and we use Monte Carlo stochastic simulation method to solve it in this section.

To improve the accuracy of the Monte Carlo algorithm, for the unknowns x_g and x_b, we obtain the range of changes of x_g and x_b in each transaction through the actual meaning behind the two unknowns and then narrow the feasible domain of the programming model. Then its accuracy is improved to a certain extent in the process of randomly generating feasible solutions. In addition, increasing the number of random numbers as much as possible can also improve the quality of the feasible solution and make the feasible solution as close to the optimal solution as possible.

After determining the investment strategy, we analyze the trend of the total assets of investors and get the following two pictures.
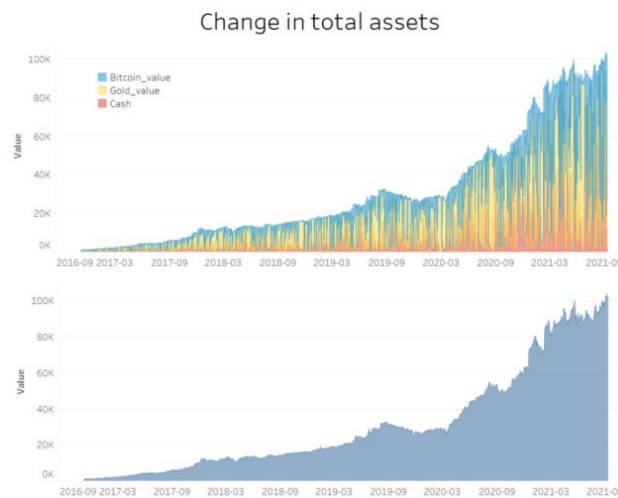

Figure 4-1: The changing trend of total assets and the area map of each part

The area chart at the top represents the time sequence chart of C, G and B assets, that is, the proportion of the three parts of the total assets every day in these five years. The larger the area of color, the greater the proportion in the total assets. From this area chart, we can see that in terms of the choice of investment strategy, because we set the investor as a stable investor, we tend to invest a large portion of our assets in cash and gold in the investment process. even if the market value of Bitcoin rises rapidly, but because of its volatility, the investor does not concentrate too much of his assets in the riskier Bitcoin. This also reflects from one aspect that the planning model we have established is in line with expectations and considers both risks and benefits.

The chart below shows the return curve generated by investors through our investment strategy, which shows that the total income curve shows a fluctuating rise, from the initial $1000 to $103635.17 after five years, with a total profit of $102635.17.

References

[1] Sun, RQ. A study of LSTM neural network-based price trend prediction model for U.S. stock indexes. Capital University of Economics and Business, 2017.

[2] Mao, JH. Study on the factors influencing the accuracy of stock market time series prediction based on LSTM deep neural network. Jinan University, 2017.

[3] Chen, L. Research on financial time series forecasting algorithm based on LSTM model. Harbin Institute of Technology, 2019.

[4] Li, Y, Design and implementation of an intelligent stock prediction system based on deep neural networks. Northwestern University, 2019.

[5] Liu, DX, Xin, M. Research on wavelet neural network prediction model for time series, 1999.